Toward Understanding Higher-Adaptive Systems

Brenda M. Ng
(925) 422-4553
ng30@llnl.gov

This project is a two-year effort that seeks to understand higher-adaptive systems, which are systems that can modify their structures and behaviors in response to attempts at detection or regulation. These systems are ubiquitous: in the real world, there are many entities, such as money launderers and cyber intruders, whose fundamental behavior changes upon probing or intervention by an observer. Such a system outputs observations (*e.g.*, an unintentional trail of evidence connected to its activities) and adversarial actions (*e.g.*, direct assaults/countermoves against its opponent). In particular, these actions can span a spectrum of aggression, from limiting information available to its opponent to misleading the opponent into making the wrong moves or assumptions.

Project Goals

The objective of this work is to explore and extend as necessary current decision-theoretical frameworks and algorithms for solving real-world adversarial problems, especially those involving adversaries that are higher-adaptive, capable of disinformation and deceptive actions. The results of this work can provide foundational knowledge for building a computationally efficient framework that can characterize and respond to dynamically changing, deceptive adversarial systems. This knowledge will be invaluable for future advanced studies of even more adaptive and aggressive adversarial systems, such as those that limit resources as well as information from their opponents. This type of study has scientific merit in both the AI and game theory communities; it also provides the basis for addressing significant national security threats.

Relevance to LLNL Mission

This project is relevant to furthering LLNL's missions in Inference and Adversarial Modeling. This work can provide important insights about real-world adversarial modeling and higher-adaptive systems, with applications in biological systems (*e.g.*, regulatory networks), law enforcement (*e.g.*, money laundering and drug trafficking), and homeland security (*e.g.*, terrorist networks, cyber attacks, and proliferation of weapons of mass destruction).

FY2009 Accomplishments and Results

In FY2009, we used money laundering as the motivating application. Money laundering is a crime in which the funds from illegal activity are disguised to appear legitimate. It is an extremely pervasive crime and can be difficult to detect, as criminals often try to diffuse the "money trail" via a complex series of financial transactions. In this work, we created a decision-theoretic model of the money laundering process, from the perspectives of both criminals and law enforcement. To model the money laundering process, we used an interactive partially observable Markov decision process (I-POMDP), which extends POMDPs for modeling multiagent adversarial systems.

In an I-POMDP, each agent maintains beliefs about the physical states of the environment, and the models of other agents (*e.g.*, how each of the other agents might perceive or act in the same environment). This makes I-POMDP

States available to the Red Team and the Blue Team in our I-POMDP model of the money laundering process. The Red Team's state denotes the location of money being laundered, while the Blue Team's state denotes the location of sensors deployed by law enforcement to surveil money laundering activities.

		Red	Blue
DP	Red's starting state: dirty pot	✓	
N	Blue's starting state: no sensors deployed		✓
BA	Domestic bank accounts	✓	✓
IN	Insurance products	✓	✓
SE	Securities	✓	✓
OF	Offshore account	✓	
SH	Shell company	✓	✓
TR	Trust	✓	✓
LO	Corporate loans	✓	✓
CA	Casino account	✓	✓
RE	Real estate	✓	✓
CP	Red's winning state: clean pot	✓	

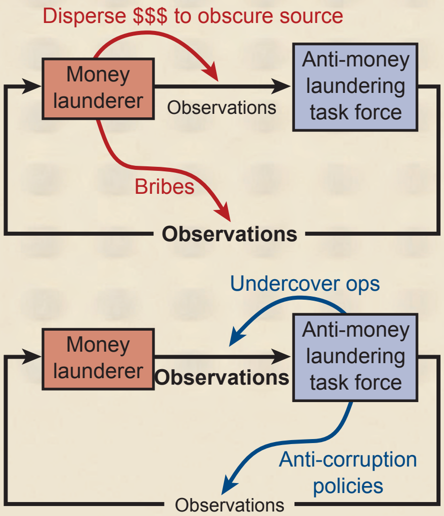

Figure 1. Schematic for understanding higher-adaptive systems. Agents can take actions to increase their own observations about the opponent, as well as actions that limit the opponent's observations.

novel with respect to other multiagent modeling frameworks in that it incorporates the notion of *nested intent* into the belief of each agent, allowing for the modeling of agents that "game" against each other, as in our money laundering scenario.

The agents in our model are the Red Team (money launderers) and the Blue Team (law enforcement). Red's goal is to evade capture while moving assets through a financial network, while Blue's goal is to find and confiscate Red's assets. The joint state consists of the locations of Red's assets and Blue's sensors. Some observations are automatically generated according to reporting requirements mandated by the Bank Secrecy Act (BSA), while other observations are obtained through actions related to active surveillance. Our I-POMDP model consisted of 99 joint states; four actions, and four observations for the Red Team; and nine actions and 11 observations for the Blue Team. Thus, the overall size of our problem (measured in terms of states × actions × observations × players) is nearly 20 times that of the largest previously solved instance in the I-POMDP literature.

To solve an I-POMDP is to determine the *optimal policy*, which for every possible observation produces an optimal action that maximizes the agent's expected reward. Previous work has shown that value iteration can be used to solve I-POMDPs. As part of our implementation of the value iteration algorithm, we applied approximations such as the *interactive particle filter* (I-PF) to address the belief complexity that increases with the number of states, and *reachability tree sampling* (RTS) to address the policy complexity that increases with number of time steps or *horizons* in the decision process. To make our problem tractable, we applied RTS to prune not only the agent's policy tree, but also the policy tree of its opposing agent. We also experimented with limited look-ahead strategies for the opposing agent.

The table and Figs. 1 and 2 illustrate our process.

Related References

1. Doshi, P., and P. Gmytrasiewicz, "Monte Carlo Sampling Methods for Approximating Interactive POMDPs," *Journal of Artificial Intelligence Research*, **34**, pp. 297–337, 2009.
2. Gmytrasiewicz, P., and P. Doshi. "A Framework for Sequential Planning in Multiagent Settings," *Journal of Artificial Intelligence Research*, **24**, pp. 49–79, 2005.
3. United States Treasury, "2007 National Money Laundering Strategy," Technical Report, Office of Terrorism and Financial Intelligence, United States Treasury, 2007.

FY2010 Proposed Work

In FY2010, we will shift our focus from passive adversaries with fixed dynamics to deceptively aggressive adversaries with adaptive dynamics. We will augment our framework to 1) include model learning capabilities for inferring the adversary's unknown dynamics; and 2) address deception and unreliability in the adversary's observable outputs.

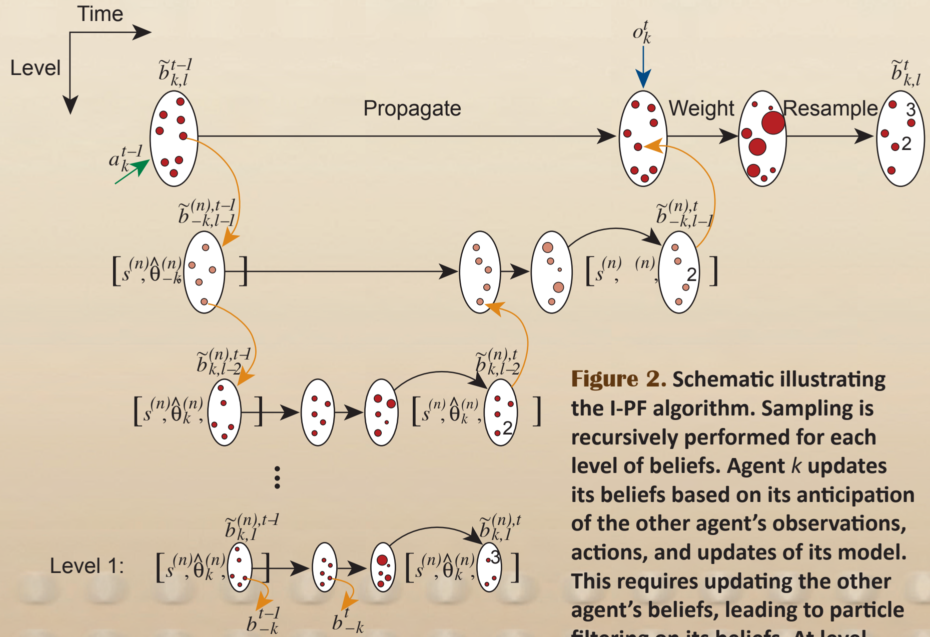

Figure 2. Schematic illustrating the I-PF algorithm. Sampling is recursively performed for each level of beliefs. Agent *k* updates its beliefs based on its anticipation of the other agent's observations, actions, and updates of its model. This requires updating the other agent's beliefs, leading to particle filtering on its beliefs. At level 1, agent *k* uses a POMDP belief update to infer the other agent's level 0 belief.